참고: LGBM, XGBOOST는 결측치 처리가 필요하지 않음

XGB -> extra boost 이게 존나 느림. 성능은 굿

LGBM -> light boost 이게 존나 빠름. 성능은 굿 but xgb만큼은 아님 -> 대회에서 디폴트로 준 값.

첫날: LGBM -> 하이퍼 파라미터 튜닝

결과 -> baseline\_withNA-RandomCV.csv

둘째날: LGBM 데이터 스케일링

결과 -> baseline\_notNA-iterative-withNorm-Random1000-test.csv

셋째날: XGB -> 하이퍼 파라미터 튜닝

결과 -> baseline\_withNA-RandomCV-XGB.csv

이번 주말: XGB의 성능을 높이자

트랙 1 -> CV 방식을 비교하자.

1. GRID CV
2. 베이지안 CV

토요일정도까지

트랙 2 -> feature selection

1. 디폴트 feature importance
2. Permutation importance

오늘 ~ 토요일

내가 할 실험: PCA 50개 피처로 해보게

50 + CV + importance

Importance 50 cv